

# *SRB Replicated Data Management for Cooperative Computing*

Arcot (Raja) Rajasekar  
San Diego Supercomputer Center  
[sekar@sdsc.edu](mailto:sekar@sdsc.edu)



SAN DIEGO SUPERCOMPUTER CENTER

*A National Laboratory for Computational Science & Engineering*

## *What is SRB?*

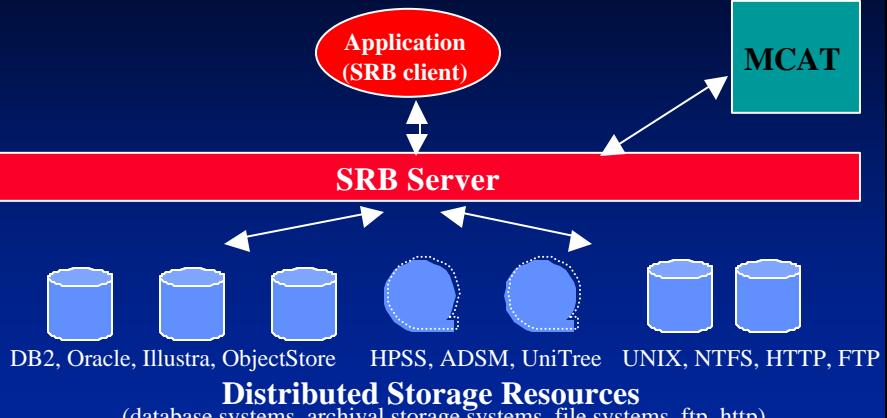
- SRB is an Intelligent Data Access System
- SRB provides federated access to datasets
- SRB provides protocol transparency to diverse and distributed storage systems
- SRB provides location transparency to distributed datasets
- SRB provides access transparency to remote user



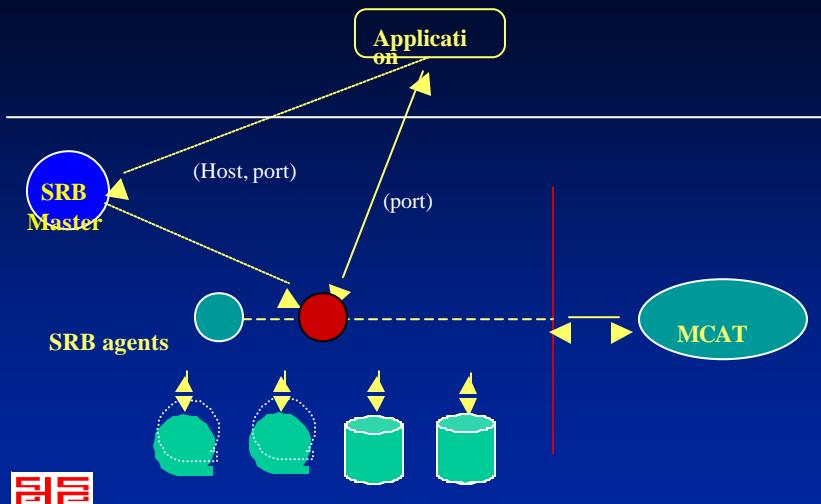
SAN DIEGO SUPERCOMPUTER CENTER

*A National Laboratory for Computational Science & Engineering*

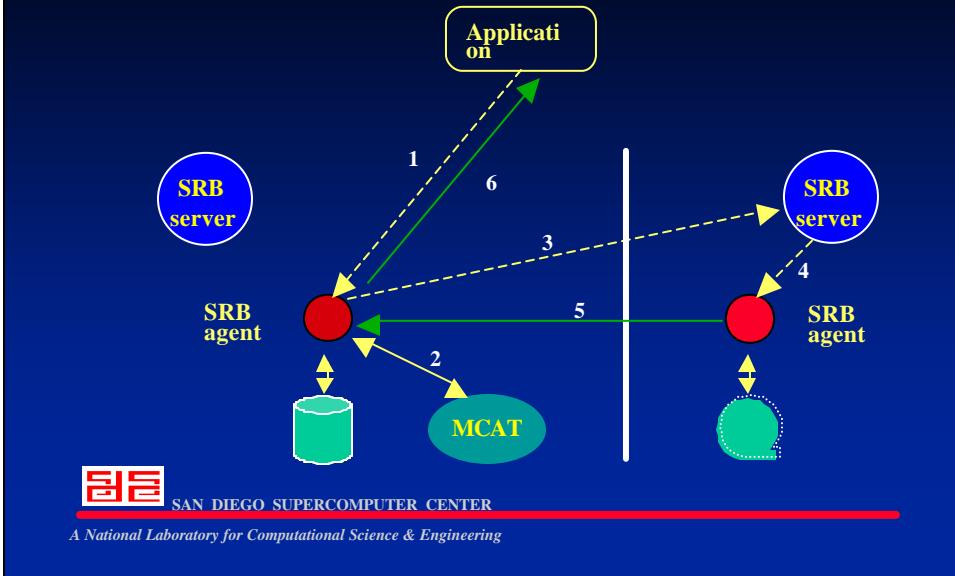
## *The Storage Resource Broker is Middleware*



## The SRB Process Model



## *Federated SRB Operation*



## ***SRB Concepts (1)***

- **Abstraction of User Space**
  - no domain dependence
  - no user accounts needed on remote servers
- **Abstraction of Resources**
  - Logical Resource Definitions - bundling
  - Resource type and Access protocol transparency
- **Abstraction of Data and Collections**
  - Persistent Identifier and Global Name Space
- **Uniform Access Methods**



SAN DIEGO SUPERCOMPUTER CENTER

A National Laboratory for Computational Science & Engineering

## ***SRB Concepts (2)***

- **Provide Scalability**

- Hosts
- Resource Types
- Resources
- Collections
- Data Objects - size and number
- Users & Groups
- Methods
- MetaData



SAN DIEGO SUPERCOMPUTER CENTER

*A National Laboratory for Computational Science & Engineering*

## ***SRB Concepts(3)***

- **Provide Logical Abstractions**

- srbSpace - an abstract storage space
- Resource Types - resource defined by properties
- Resources - resource identified by name and type
  - multiple resources tied together as a single resource
- Collections - abstraction over directory structure
  - distributed & curated
- Datasets - identified by properties
- Users - authenticated across hosts/networks
- Domain - abstraction over physical domains
- Metadata Schema/Attributes



SAN DIEGO SUPERCOMPUTER CENTER

*A National Laboratory for Computational Science & Engineering*

## ***SRB Concepts(4)***

- Replication of Datasets
- Collections for logical co-location
- Containers for physical co-location
- Access Control Lists for Authorization
- Ticket-based Access
- Auditing
- Authentication and Encryption (SEA)
- Server-side proxy Operations
- Metadata-based Discovery
- Rich Interface - programmatic & interactive



SAN DIEGO SUPERCOMPUTER CENTER

A National Laboratory for Computational Science & Engineering

## ***SRB Space***

DR - Data Repository  
DL - Dig Library  
MC - Meta Catalog  
CP - Comp Process/  
SRB Client

A National Laboratory for Computational Science & Engineering

## *MCAT: Metadata Catalog*

- Stores metadata about
  - Data sets, Users, Resources, Proxy Methods,
- Maintains replica information for data & containers
- Provides “Collection” abstraction for data
- Provides “Global User” name space & authentication
- Provides Authorization through ACL & tickets
  - data, collection, resources and methods
- Maintains audit trail on data & collections
- Maintains metadata for methods and resources
- Provides Resource Transparency - logical resources
- Implemented as a relational database - Oracle or DB2



SAN DIEGO SUPERCOMPUTER CENTER

A National Laboratory for Computational Science & Engineering

## *SRB Replication Concepts*

- Replication is a core functionality in SRB
- Global Name Space (hierarchical)
  - local name independence
  - replica can reside in any type of resource
- Persistent Id: data movement independence
- Access Control at Replica Level
- Resource Access Control
- Replicas created using SRB or from outside
- Semantic Replicas & Syntactic Replicas
- Typing of Replicas: Archive, Cache, Temporary



SAN DIEGO SUPERCOMPUTER CENTER

A National Laboratory for Computational Science & Engineering

## ***SRB Data Replication Support***

- **Synchronous Replication**
  - Replication via *Logical Resource* definition
  - integrated into *open/create & write* function
  - Can choose:  $k$  out of  $n$
  - Associate replication with containers/collections
  - Consistency
- **Asynchronous Replication - Offline**
  - `srObjReplicate API`, `Sreplicate command`, GUI
- **Out of Band Replication - outside SRB**
  - Registering of Replicas using `srRegisterReplica API`



SAN DIEGO SUPERCOMPUTER CENTER

A National Laboratory for Computational Science & Engineering

## ***SRB Data Replication Support***

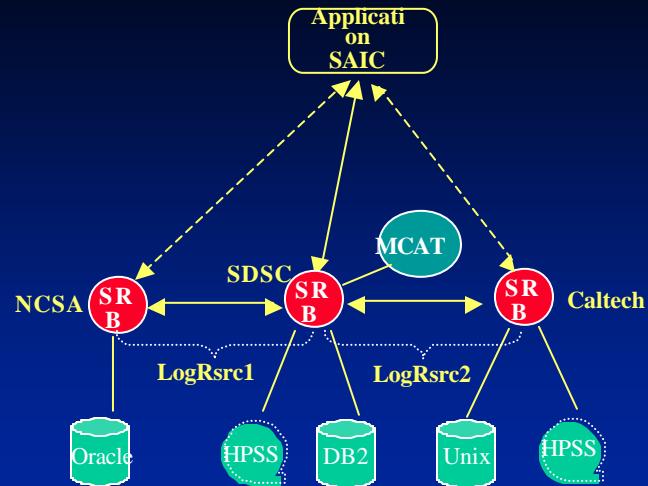
- **Choice at Read**
  - any replica
  - specific replica (by copy number)
  - round-robin
  - “nearest”
  - by resource characteristics
  - by timestamp or other characteristics
  - data itself may be identified by meta characteristics
    - user defined metadata & annotations
    - data type, owner, comments, ...



SAN DIEGO SUPERCOMPUTER CENTER

A National Laboratory for Computational Science & Engineering

## *Data Replication*



A National Laboratory for Computational Science & Engineering

## *SRB API*

- **Programmatic API**
  - High-level API
  - Low-level API
  - SRB Manager API
- **Command Level Interface - Scommands**
- **Graphical User Interface - Java Browser, NT Browser**
- **Web Utilities**
- **Transparent Access**

SAN DIEGO SUPERCOMPUTER CENTER

A National Laboratory for Computational Science & Engineering

## *High & Low-level API*

- **Low-level API**
  - talks to resource drivers
  - no registration of data sets in MCAT
  - no authentication through MCAT
  - User provides all information
- **High-level API**
  - Uses low-level API to access resources
  - Registers data management information in MCAT
  - Uses MCAT for authentication and meta information
  - Uses MCAT for resource and data discovery
  - Access/store data in remote SRB



SAN DIEGO SUPERCOMPUTER CENTER

A National Laboratory for Computational Science & Engineering

## *Low-level API*

- `srpNetOpen(conn, storType, host, fileName, mode)`
- `srpNetCreate(conn, storType, host, fileName, mode)`
- `srpNetClose(conn, fd)`
- `srpNetUnlink(conn, storType, host, fileName)`
- `srpNetRead(conn, fd, buffer, length)`
- `srpNetWrite(conn, fd, buffer, length)`
- `srpNetSeek(conn, fd, offset, whence)`
- `srpNetSync(conn, fd)`
- `srpNetStat(conn, storType, host, fileName, statBuf)`
- `srpNetMkdir(conn, storType, host, dirName, mode)`
- `srpNetRmdir(conn, storType, host, dirName, mode)`
- `srpNetChmod(conn, storType, host, fileName, mode)`



SAN DIEGO SUPERCOMPUTER CENTER

A National Laboratory for Computational Science & Engineering

## *High-level API*

- `srObjOpen(conn, objChar, mode, collectionName)`
- `srObjCreate(conn, objName, objType, resourceName, collectionName, pathName, size)`
- `srObjClose(conn, od)`
- `srObjUnlink(conn, objChar, collectionName)`
- `srObjRead(conn, od, buffer, length)`
- `srObjWrite(conn, od, buffer, length)`
- `srObjSeek(conn, od, offset, whence)`
- `srObjMove(conn, objChar, collectionName, newResourceName, newPathName)`
- `srObjReplicate(conn, objChar, collectionName, newResourceName, newPathName)`
- `srObjProxyOpr(conn, Operation, sourceDesc, targetDesc)`
- `srRegisterReplica(conn, objChar, collectionName, newResourceName, newPathName)`



SAN DIEGO SUPERCOMPUTER CENTER

A National Laboratory for Computational Science & Engineering

## *High-Level API (contd ...)*

- `srGetDatasetInfo(conn, objChar, collectionName, resultStruct, requiredNumber)`
- `srGetMoreInfo(resDesc, resultStruct, requiredNumber)`
- `srGetDataDirInfo(conn, conditionList, selectList, resultStruct)`
- `srModifyDataset(conn, objId, collectionName, newValue1, newValue2, modifyType, resourceName, pathName)`
- `srCreateCollect(conn, parentCollectionName, childCollectionName)`
- `srListCollect(conn, CollectionName, flag, resultStruct)`
- `srModifyCollect(conn, CollectionName, newValue1, newValue2, newValue2, modifyType)`
- `srModifyUser(conn, newValue1, newValue2, modifyType)`
- `srSetAuditTrail(conn, setValue)`



SAN DIEGO SUPERCOMPUTER CENTER

A National Laboratory for Computational Science & Engineering

## *Scommands*

- Sinit - initialize S-environment
- Sextit - clean up
- Sman - get manpage for Scommand
- Scat - display srableObject on screen
- Sput - copy local file into srbSpace
- Sget - copy srableObject to local space
- Sappend - append to srableObject
- Srename - change srableObject name
- Srm - remove srableObject
- Schmod - change/grant access to srableObject
- Scd - change collection
- Spwd - display current collection
- Sls - list collection
- Smkdir - make new collection
- Srmdir - remove old collection
- SgetD - get srbObject information
- SgetR - get resource information
- SgetU - get user information
- SmodD - modify srbObject info
- SmodU - modify user info
- Stoken - get native type information
- Scopy - copy srableObject in another collection and under another name
- Sreplicate - clone object in new resource - same internal id
- Smove - move srableObject to new collection or resource



SAN DIEGO SUPERCOMPUTER CENTER

A National Laboratory for Computational Science & Engineering

## *Scommands (contd ...)*

- ingestUser - adding a new user or group
- ingestResource - adding a new resource
- ingestLogicalResource - making a new resource grouping
- addLogicalResource - adding to a resource grouping
- ingestLocation - adding new location information
- ingestToken - adding new native types  
(eg. resourceType, objectType, userType, domainName, ActionType, ...)



SAN DIEGO SUPERCOMPUTER CENTER

A National Laboratory for Computational Science & Engineering

## *Web Utilities*

- Sgetw - copies a SRBobject into server site
- Sputw - copies local file in SRBspace
- Scatw - displays SRBobject on browser (handles types)
- Slsw - displays information of SRBobjects



SAN DIEGO SUPERCOMPUTER CENTER

*A National Laboratory for Computational Science & Engineering*